DOCUMENT RESUME

ED 088 954 TM 003 522

AUTHOR Sherman, Thomas M.; Winstead, John C.
TITLE A Formative Approach To Student Evaluation Of Instruction.
PUB DATE Apr 74
NOTE 15p.; Paper presented at the annual meeting of the American Educational Research Association (Chicago, Illinois, April 15-19, 1974)

EDRS PRICE MF-$0.75 HC-$1.50
DESCRIPTORS College Instruction; College Students; *Course Evaluation; Effective Teaching; Feedback; *Formative Evaluation; *Instructional Improvement; Rating Scales; *Teacher Rating; Teaching Quality; Teaching Skills

ABSTRACT

A procedure for using the concept of formative evaluation in student evaluation of instruction is presented. Sixty undergraduate students in an educational psychology course completed daily evaluations of every instructional program. The procedure was carried out quickly and efficiently through the use of a 5 point scale and a standard reporting format. The results were reviewed daily by the instructor. Results indicated the procedure is both reliable and valid. In addition, the procedure appears to offer data useful in revising instructional procedures during the conduct of the course. Thus, the improvement of instruction is greatly facilitated due to the specificity of the data produced. Other implications for improving the quality of instruction are discussed relative to meeting individual student needs. A daily record sheet and the rating scale are in the document. (Author/BB)

ED 088954

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A Formative Approach to Student Evaluation of Instruction

by

Thomas M. Sherman
and
John C. Winstead

Virginia Polytechnic Institute
and State University
College of Education
Blacksburg, Virginia

Paper presented at the 1974 annual meeting of the American Educational Research Association, Chicago, Illinois

TM 003 522

Student evaluation of instruction has been a much discussed issue resulting in diverse and often contradictory findings. Several issues seem to comprise the majority of concern with student ratings; among these are: the reliability of such ratings, the validity of the ratings, the use of the ratings and the effect of student ratings on teacher behavior change. Studies have been published which point to students' ability to reliably rate instructors (e.g. Costin, 1968) and other studies verify the validity of these ratings (e.g. McKeachie, 1969). The more common proposed uses for student ratings include providing course end feedback for instructors, evaluating teaching competence for promotion purposes and salary adjustments and, providing the student body with information for selecting courses (or more properly, instructors). The final issue, the effect of student ratings on instructor behavior, (i.e., improving the quality of teaching) has not received a great deal of attention. What evidence does exist is contradictory and covers the total spectrum of helpful to harmful. Tuckman and Oliver (1968) found feedback from student evaluations helpful in producing student achievement gains when feedback was given to high school instructors on a mid-semester student evaluation. Miller (1971), however, found no difference in student ratings of instruction between a mid-semester and end-of-semester evaluation when freshmen teaching assistants were given feedback at the mid-semester. Finally Oles and Lencoski (1973) found that graduate teaching faculty did change their self-ratings as a consequence of receiving student evaluations but that some of the instructors made changes in "a direction exactly opposite of what would have been expected from the student feedback information." (Oles and Lencoski, 1973)

Considering the importance student evaluations are presently being given it is puzzling that few investigations have been conducted on the effect of student evaluation feedback on instructor teaching behavior. However, this research gap is not too surprising for several reasons. Generally, instructors do not appear to place a great deal of faith in course-end ratings, this apparent lack of faith mitigates against their use in a positive manner to promote instructional changes. Also, traditionally, the teaching function has been viewed as the individual perogative of each instructor. The notion of academic freedom often precludes criticism of instructional procedures and professional ethics often dictate that such criticism should not be publically voiced. Perhaps, the most compelling reason behind the lack of research concerning teacher change as a result of student evaluation lies in the nature of the evaluation instruments used for student ratings. The type of student evaluations generally used provide almost no information with enough utility to suggest teacher change.

The procedure usually followed in gathering student evaluation of instruction is to give students a standard or quasi-standard questionnaire at or near the completion of a course. This questionnaire usually contains items such as "This course was poorly organized." (Finkbeiner, et. al., 1973), "Now that you've taken this course, would you recommend it to a friend who did not have to take it?" (Hartley and Hogan, 1972) and "Achievement of course objectives was helped by class activities." (College of Education Student Evaluation Form, 1973). Students are asked to rate the instructor on each item, using a scale, usually from 1 to 5. Instructor feedback consists of an average score for each item.

There are three major difficulties with this approach. First, course-end evaluations contribute very little to the improvement in quality of either teaching or learning because of the timing of the information gained. The

course is over when the evaluation data is received. Second, the type of information gained is ordinarily so general and non-specific that it provides no reasonable specific goals for instructor behavior change. The third major problem lies in the nature of what is evaluated. As Hartley and Hogan (1972) noted, most course evaluations do not mention the students' self-development as a result of the course, thus, ignoring student estimates of outcome or personal growth.

Many of the problems resulting from the use of traditional forms of student evaluation of instruction may be dealt with if the distinction between formative and summative evaluation is employed in seeking a solution. Bloom, Hastings and Madaus (1972) distinguished between formative and summative evaluation based upon the purpose, portion of the course covered (time) and level of generalization to be included in the evaluation. Formative evaluation is defined as being made "...during a course, when (presumably) changes can be made in the transactions of subsequent instruction on the basis of current attainment (Bloom, Hastings and Madaus, 1972, p. 262)." Summative evaluation is defined by Bloom, et. al. (1972) as '...used to designate student assessment at the end of a course or topic unit, that is, when no subsequent changes in treatment for that learning will be made (p. 262)." While this distinction has not apparently been applied to student evaluation of instruction, it appears to hold promise in solving some of the problems noted by others with student evaluation of instruction. This paper reports a formative system for student evaluation of university instruction, the analysis of the data and uses of the data to improve instruction. The formative evaluation procedure reported in this paper included two dimensions often included in student evaluations: quality of instruction and student feeling of personal gain.

Procedures

Subjects and Setting

Sixty students enrolled in a teacher education program served as subjects while taking a required course in educational psychology. The class met thirty times for fifty minutes during the spring quarter. On the first day of class the students were given the following materials and instructions:

1. A course proclamation which outlined the general objectives for the course, the class procedures and routines to be followed and the criteria to be used in grading.

2. A class daily record sheet (see figure 1) on which students recorded formative evaluations of the instructor and themselves and the number of daily behavioral objectives required and completed on a day to day basis.

3. An explanation (see figure 2) detailing the use of the class daily record sheet and formative evaluation procedures.

4. A standard size file folder in which all evidence of completion of daily objectives was kept as well as the class daily record sheet.

Students were divided into groups of five to seven members. A group folder was distributed to each group for assignments that were to be completed by the groups as a whole. The students were requested to put all the folders of students in the group together and with the group folder and to put a rubber band around all the folders. This procedure facilitated the daily distribution and collection of student folders. All folders were checked following every class; this allowed the students' daily evaluations of instructional activities to be reviewed.

Formative Evaluation

Formative evaluation was carried out quickly and efficiently through the use of a five point scale similar to the type scale often used with course evaluation (see figure 2). Students were given several verbal examples of

possible class experiences for rating the class on a daily basis and asked to use what they considered their "average" experience as a reference point for the rating scale. Two categories were rated each day; the student's evaluation of the instructional presentation and the student's evaluation of his personal gain from that day's presentation. Students were asked to become familiar with the rating scale and were reminded to complete the evaluations during the last minute of the first three classes. Following this, no reminders were issued and the procedure generally appeared to take no longer than 10 to 30 seconds for any student once the routine was established.

At the completion of the course students were asked to total the number of objectives completed and required. The evaluations under self and instructor were also totaled and averaged yielding an overall course evaluation score.

Results and Discussion

Several factors were considered in evaluating the results of the formative evaluation data. First, reliability results indicated that students were able to rate on a day to day basis with a high degree of reliability (KR_{20} = .882). In addition, the correlation between student evaluations of the instructor and their evaluations of personal value of the day's activities was .745. These procedures were carried out following the completion of the course. While the overall ratings by students were quite high (average personal evaluation, 1.81, average instructor evaluation, 1.92) and indicated a high level of satisfaction by the students, this data yielded no significant information relative to a particular lesson. However, using the reliability data, it was discovered one day in particular did not contribute to high reliability. A review of the instructor's comments which were recorded following each day's lesson indicated he felt the lesson for this day to be in need of revision (e.g., "poor objectives," "poor lecture presentation"). In this case the reliability data substantiated

the instructor's feelings about that day's class. This indicated a need for revision of this lesson prior to re-presentation.

Perhaps of even more importance was the role the daily evaluations played in solving instructional problems with individual students and particular concepts during the course. A review of the daily ratings by the instructor on a class to class basis revealed that two students were consistently rating the classes as being of little value to them and of low instructional quality. These students were given the opportunity to meet course objectives in a manner more consistent with their personal learning styles which resulted in increased feelings by these students regarding the value and quality of the course. In addition, when there appeared to be a great deal of variation in student ratings for any single day, this was interpreted as indicating a lesson which lacked clarity. Hence, this served as a cue to the instructor for some needed remedial work to be offered regarding the concepts, skills or values taught for that day.

Perhaps the major reason previous use of student evaluation of instruction has not resulted in instructor change (e.g. the same instructors seem to get low ratings quarter after quarter) is that these course evaluations have not been designed to facilitate instructor change. Rather, most of these instruments appear to be designed to give a global overview of how an instructor functions. This type of summative evaluation is useful when administrative decisions regarding salary adjustment are to be made or a student wishes to know who is considered a "good teacher." But, when the goal of the instrument is to foster improved instruction, evaluation after the fact in terms of generalities is not helpful because instruction consists of specific behaviors.

Whatever classroom interaction either helped or hindered the classroom learning situation can seldom be appropriately manipulated with summative

evaluation only. Verbal and nonverbal teaching habits (eye contact versus looking out the window or at the ceiling, throat clearing, OK's, flat expressionless speaking, unthinking animation, poor teaching materials, etc.) are unlikely to change as a result of the global non-helpful statements such as "instructor's personality was a negative factor in this course" usually found on summative course evaluations. Even when factors are identified to explain the dimensionality of an instrument (e.g. Finkbeiner, et. al., 1973) it is doubtful that scores reported on any dimension such as an instructor characteristics factor will provide information specific enough to suggest some positive instructor action. Formative evaluation carried out on a class session by class session basis has the potential advantage of behavioral specificity. This specificity can facilitate changes in instructional programs because of the immediacy and specific nature of the feedback received from students. With formative student evaluation, student feedback can be directed at specific teaching techniques and specific lessons as well as include student evaluation of personal gain (outcome) for each lesson. In this manner, the relative contribution of each instructional period to an overall evaluation of instruction may be assessed.

It is as unrealistic to expect students to be uniform as it is to expect teachers not to vary. Teaching styles differ greatly and so do learning styles. Higher education has, to an unacceptable degree, ignored individual student differences. The major mode of instruction is the lecture; however, for many students this is not the optimal learning style though it probably serves well for the vast majority. Most students who have different learning styles have been removed from the educational system by the time they reach college. Still, a few do remain who will profit greatly from a different instructional mode. Individuals who are not or who feel they are not responding satisfactorily to the dominant mode of instruction may be identified through their daily ratings.

As more and more pressure is put on public institutions of post-secondary education to accept a broader range of students, and to some extent guarantee these students success, it will undoubtedly become more important to have effective ways of identifying individual students who are not responding well. In order to be useful and helpful to the student, this must be done early and continually in the instructional program.

Some light may be shed on two controversial issues regarding student rating data from the results of the reported formative evaluation. First, the issue raised by Hartley and Hogan (1972) that "we should be at least equally interested in the students' judgment of their own self-development as affected by the course" (p. 249) as in the students' subjective judgment of an instructor's behavior may be superfluous. The high correlation between student daily ratings of the instructor and student evaluation of personal gain suggests either may serve as an indicator of the other. However, this may be true only when the ratings are made on a day to day basis. Second, Rodin and Rodin (1972) indicated that students fail to recognize good teaching when they are exposed to it and Kosoff (1971) has expressed doubt whether good teaching can be judged at all. While the data do not provide a conclusive answer as to whether students can identify good teaching, there is evidence to indicate they can. The one day which was discovered to be a low contributor to the reliability coefficient was also felt to be a poor day by the instructor, supposedly an expert in designing and evaluating instructional programs. This data at least provides some notion of validity, even if only based on the average perception by the class and the perception of the instructor. In any event, where such agreement exists regarding the ineffectiveness of a particular lesson, there is compelling reason for changes to be made in the instructional program for that day.

Few people are naturally good teachers; students often appear to feel that universities get less than their share of the few who are naturally endowed. Even those who do possess quality instructional skills cannot expect these skills to be maximally effective for all students. If teachers are willing to seek out the students with whom they are ineffective and to also meet the needs of these individuals, the probability of significant learning occurring will increase greatly. The process of quality instruction is indicated as the probability of learning increases. The formative student instructional evaluation system described here is designed to enhance this process.

Bibliography

Bloom, B.S., Hastings, J.T. and Madaus, G.F. Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill, 1971.

College of Education Student Evaluation Form, College of Education, Virginia Polytechnic Institute and State University, 1973.

Costin, F. A graduate course in the teaching of psychology: description and evaluation. Journal of Teacher Education, 1968, 19, p. 425-432.

Finkbeiner, C.T., Luthrop, J.S. and Schuerger, J.M. Course and instructor evaluation: some dimensions of a questionnaire. Journal of Educational Psychology, 1973, 64, p. 159-163.

Hartley, E.L. and Hogan, T.P. Some additional factors in student evaluation of courses. American Educational Research Journal, 1972, 9, p. 241-250.

Kossoff, Evelyn. Evaluating college professors by "scientific" methods. The American Scholar, 1971-72, 41, p. 79-93.

McKeachie, W.J. Teaching Tips. A guidebook for the beginning college teacher. (6th ed.) Lexington, Mass.: D.C. Health, 1969.

Miller, M.T. Instructor attitudes toward, and their use of student ratings of teachers. Journal of Educational Psychology, 1971, 62, p. 235-239.

Oles, H.J. and Lencoski, A. Changes in an instructor's self-rating resulting from feedback from student evaluations. Selected Documents in Psychology. Journal Supplement Abstract Service, MS, No. 309, 1973.

Rodin, M. and Rodin, B. Student evaluations of teachers. Science, 1972, 177, p. 1164-1166.

Tuckman, B.W. and Oliver, W.F. Effectiveness of feedback to teachers as a function of source. Journal of Educational Psychology, 1968, 59, p. 297-301.

Figure 1. Class daily record sheet used to record the number of objectives assigned and completed and formative evaluation of instructor and student gain on a daily basis

Name _____ Class Time _____

Course _____ Quarter _____

DATE	OBJECTIVES		EVALUATION	
	Assigned	Completed	Personal	Instructor

Comments:

_____

_____

_____

_____

_____

_____

_____

_____

Figure 2. Instructions given to students explaining the use of the Class Record Sheet and the Formative evaluation procedures.

Every day you are present in class you should complete your Class Record Sheet. This should remain current at all times. The Class Record Sheet is to be used to record daily progress in attempting and completing the daily objectives of the course. Each week you will be given a schedule of upcoming topics to be covered during the week.

The following procedures are to be used in recording your progress on the Class Record Sheet.

1. Write in the appropriate date in the space provided.

2. (a) The number of objectives assigned for that day and (b) the number that you have completed. You must complete 90% of the assigned objectives to be eligible to receive a grade in this course.

3. Under the heading evaluation record (a) your personal evaluation of the day's activities and (b) your evaluation of your instructor for that day.

   a) Personal evaluation of the day's activities.
   Use the following scales:

   1 = I feel this was an excellent day. I "learned" material which seems to be highly relevant to good teaching. I was very involved.

   2 = I feel this was a very good day. I "learned" material which appears to be relevant to good teaching. I was involved most of the time.

   3 = I feel this was a good day. I "learned" seemingly relevant material. I was involved a fair amount of time.

   4 = I feel this was a poor day. I "learned" some relevant material but not much. I was not involved to any great extent.

   5 = I feel this was a horrible day. I "learned' little if any material relevant to good teaching. I was almost totally uninvolved.

   This scale should be used with your past academic experience as your reference point. That is, you should consider your acquisition of knowledge and skill in this class in relation to other classes you have had. Therefore, in relation to other classes in your past experience, you may feel this class for the particular day is superior; if this were the case you would rate the class that day a one. On the other hand, you may feel a particular class is really rotten; you would rate the class as a five for that day.

b) Under the heading Instructor you should record your evaluation of your instructor for that day. This rating should be made on the basis of the clarity and rationale of the class presentations, and the quality of the instruction program you are presented. The following scale should be used:

1 = Excellent presentation, clear, concise; excellent skills programs

2 = Very good presentation and/or skills exercise. Generally very clear and understandable.

3 = Good presentation and/or skills exercise. Fairly clear and understandable.

4 = Poor presentation and/or skills exercises. Unclear and only comprehensible with great effort.

5 = Horrible presentation and/or skills exercises. Obscure, obviously irrelevant, pointless, a total waste of time.

Finally, in the space provided at the bottom of the Class Record Sheet, make any comments you feel particularly relevant for that day's class. You may not want to include a comment each day but when you do write a comment please date it. You may make any comment you wish about the quality of the instruction and it would be most helpful if these comments pointed out specific particulary positive or negative instructor behaviors which contributed to the quality of the day's class. Examples of such comments might be: "You talked too much in a monotone voice." or "Your use of examples was very poor and confusing." or "Your use of humor during the lecture was excellent and maintained interest." or "The personal attention you give each student helps in learning the material."

A Formative Approach to Student Evaluation of Instruction

Thomas M. Sherman
and
John C. Winstead

ABSTRACT

A procedure for using the concept of formative evaluation in student evaluation of instruction is presented. Sixty undergraduate students in an educational psychology course completed daily evaluations of every instructional program. The procedure was carried out quickly and efficiently through the use of a 5 point scale and a standard reporting format. The results were reviewed daily by the instructor. Results indicated the procedure is both reliable and valid. In addition, the procedure appears to offer data useful in revising instructional procedures during the conduct of the course. Thus, the improvement of instruction is greatly facilitated due to the specificity of the data produced. Other implications for improving the quality of instruction are discussed relative to meeting individual student needs.